Medicinal Chemistry
- an introduction

Oscar Belda

Principal Scientist, Project Manager

Medivir AB

oscar.belda@medivir.com

Contents

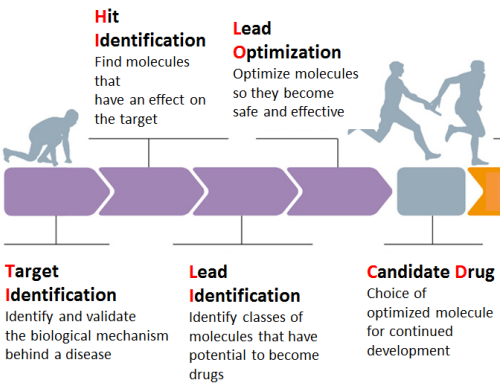
1. Introduction to Medivir AB
2. The Discovery Process – TI, HI, LI, LO and the medicinal chemist's role

Introduction to Medivir AB

- R&D dedicated company focused on **oncology**
- Spinout from Astra's antiviral research unit, currently ~74 employees
- Expertise in **polymerase** and **protease** drug targets
- 2 products from benchtop to market: **Olysio** (HCV, J&J) and **Zoviduo** (cold sores, GSK)

MEDIVIR

The Discovery Process and the Medicinal Chemist's Role



Optimize the properties of an active compound so that it succeeds in the clinical development phase

Hit Identification

Definition of Tems

- **Hit**

  Chemical substance that interacts with a target

- **Assay**

  Any combination of targets and compounds which is exposed to a detection device to measure chemical or biological activity.

- **Affinity**

  Of a ligand is its ability to bind to a biological target.

  $T + L \rightleftarrows C$; $K\downarrow d = [T][L]/[C]$ *lower* $K\downarrow d$ *then higher affinity*

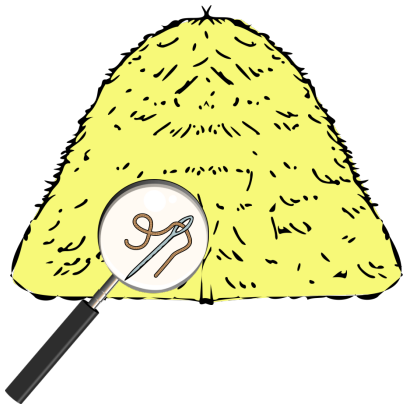
- **Druggability**

  The ability of a ligand to bind to a target with properties consistent with oral activity and adequate pharmacokinetics and pharmacodynamics

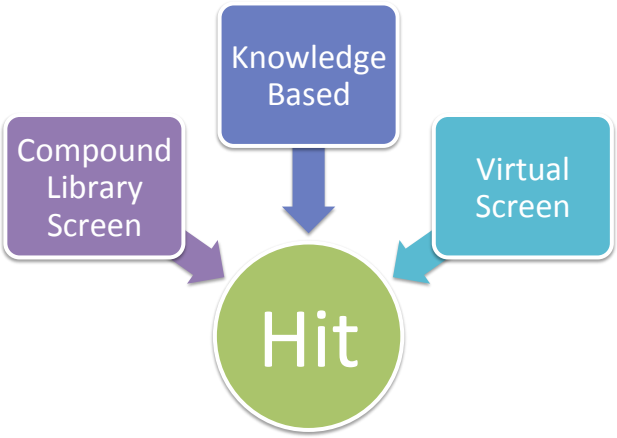
Hit Identification

Strategies

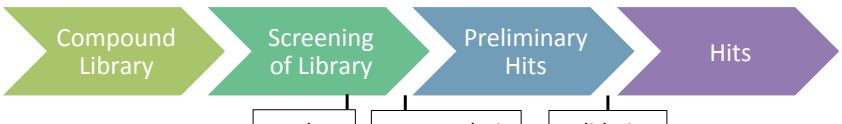
How do we find chemical compounds that have affinity for the target of interest?

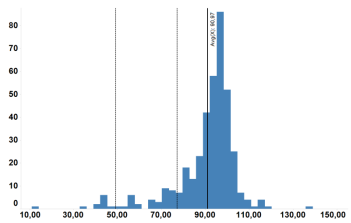


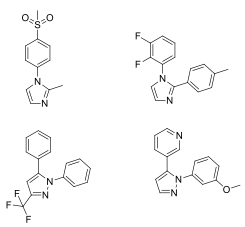
Hit Identification
Strategies



Hit Identification
Compound Library Screen







Hit Identification

Compound Library Screen

❑ **Compound Libraries**

- Chemical Diversity (chemical space) vs Target Directed
- DNA encoded Libraries

Molecular Shape Diversity
Principal Moment of Inertia (PMI)

Hit Identification

Compound Library Screen

❑ **Screen for Phenotype (-1970s)**

- Looks for a specific effect not for a specific target interaction
- High Content Screen

Hit Identification

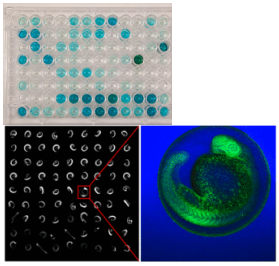
Compound Library Screen

❑ **Screen for Mechanism (1990s-)**

- Look for compounds that have affinity for a specific target
  - High Throughput Screening (HTS)
  - DNA Encoded Library (DEL) Screen
  - Fragment Based Lead Discovery (FBLD)

Hit Identification

Compound Library Screen:
High Throughput Screen (HTS)



❑ **Many types of readouts:**

- Fluorescence/Luminescence
  - Fluorescence Polarization (FP)
  - Flurorescence Resonance Energy Transfer (FRET)
- Gene reported assays
  - Luciferase
- High content imaging

❑ **HTS Triage (Data Analysis): process to separate True from False**

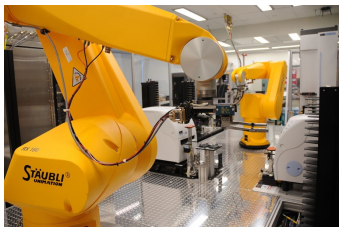
- HTS true positive hit rates 0.1-2%
- False positive: data point that appear active because of the ”noise” in the assay
  - Random: related to the biology screen
  - Non random: chemistry related
- False negatives: interference with assay, solubility

Hit Identification

Compound Library Screen:
High Throughput Screen (HTS)

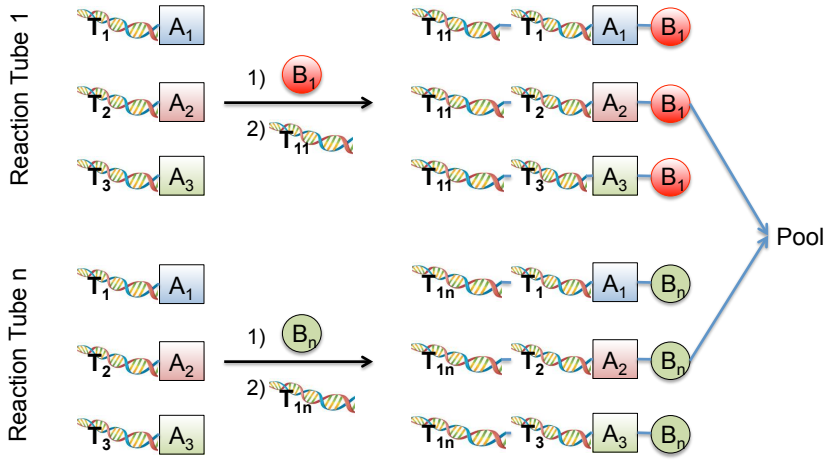
❑ **Limitations of HTS:**

- Compound libraries
  - Diversity
  - Exclusivity
  - Quality: purity of compounds, drug like properties
- Low hit rates
- High cost: library + assay + data mining + maintenance
- Difficult to measure low affinity compounds



Hit Identification

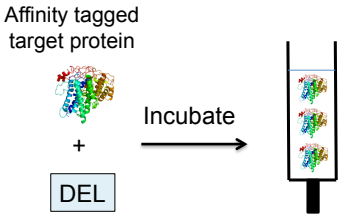
Compound Library Screen:
DNA Encoded Library (DEL) Screen



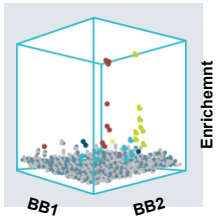
Very large libraries (billion of compounds) can be made in this way

Hit Identification

Compound Library Screen:
DNA Encoded Library (DEL) Screen



1. Wash to remove non binder compounds (enrichment)
2. Elute binders
3. Sequence DNA of binders



Families of compunds are colored

Hit Identification

Compound Library Screen:
Fragment Screening

❑ **Screening based on biophysical techniques:**

- Protein X-Ray Crystallography
- NMR
- Surface Plasmon Resonance (SPR)
- MicroScale Thermophoresis (MST)
- And many more

❑ **Advantages over HTS:**

- Possible to detect weak affinity ligands (mM range)
- The hits are usually very simple but with high binding efficiency per heavy atom

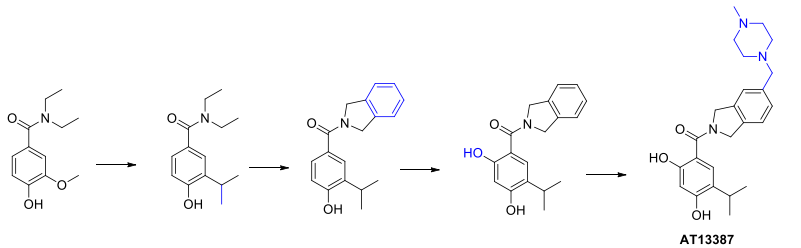
$$LE = \Delta G / HAC \quad \text{approximates to} \quad -1.37 \cdot \log(Kd) / HAC$$

HAC = number of non H atoms

- Steep SAR: from mM to nM in few steps

Hit Identification

Compound Library Screen:
Fragment Screening



Astex AT13387 Hsp90 Inhibitor

Murray et al J Med Chem 2010, 5942 and Woodhead et al J Med Chem 2010, 5956

Hit Identification

Compound Library Screen:
Fragment Screening

❑ **Limitations of Fragment Screening:**

- Relatively low throughput
- Biophysical technique specialist
- Usually needs high amounts of protein
- High concentration of ligand = high solubility requirement

Hit Identification

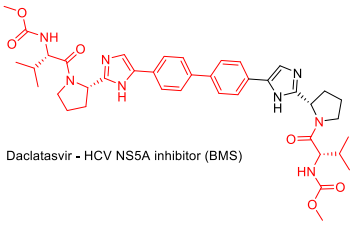
Knowledge Based

- **Modifications of natural ligands, e.g. Peptidomimetics**

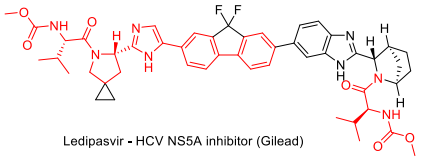
Thymidine - DNA nucleoside

Azidothymidine - HIV reverse transcription inhibitor

- **Leads from another in-house program**
- **Patent Busting (chemical evolution)**



Daclatasvir - HCV NS5A inhibitor (BMS)



Ledipasvir - HCV NS5A inhibitor (Gilead)

Hit Identification

Virtual Screen

- **Ligand Based:**

Structural Diverse ligands known to bind to the target are used

- Pharmacophore model
- Relate substructural features to affinity
- Similarity analysis

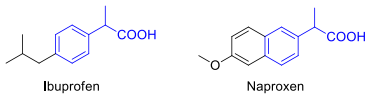
- **Structure Based:**

Docking of candidate ligands into target followed by scoring function

Hit Identification

The Medicinal Chemist's Role

- **(Structural) Pattern recognition**



Ibuprofen Naproxen

COX inhibitors

- **Design of screening libraries**

  Prior biomedical knowledge vs maximum chemical diversity

- **Potential problems of a hit and solutions**

- **Hit confirmation**

  Confirm structure, resynthesize and SAR pattern

Lead Identification

Definition of terms

- **Confirmed Hit**

  A chemical entity having a defined structure which has been shown to have a desired effect in a biological assay

- **Lead**

  A chemical entity with a defined effect in a biological assay which by subsequent modification can be transformed predictably into a clinically useful drug

- **SAR (Structure – Activity – Relationship)**

  Correlation within a chemical series on how a change in structure of a molecule affects the biological activity

Lead Identification

The perfect Lead

- Mean MW 350 (Lead-like)
- Lipinski score <= 2 (Lipinski rule of 5: LogP <5, MW <500, HBD <5, HBA <10)
- Rotatable bonds < 10
- Heavy atoms < 30
- Solubility > 50 μM
- No reactive groups
- Chemically a good "starting point"

Lead Identification

Next steps for the Medicinal Chemist

- ❑ **Is there Structure Activity Relationships (SAR) ?**
- ❑ **Learn about the liabilities and attributes of a chemical series (Profiling)**
  - Potency against target and selectivity
  - Aqueous solubility
  - Protein binding
  - Permeability
  - Metabolic stability
  - Toxicity and off target screen
- ❑ **Check IP space**
- ❑ **Check Synthesis**

The Goal is to find 3-5 independent substance classes for Lead Optimization

Lead Identification

Choosing a Lead: which one will you choose?

Compound 1	Compound 2	Compound 3
Ki = 50 μM	Ki = 0.8 μM	Ki = 5 μM
MW = 410	MW = 350	MW = 250
Papp – very low	Papp – high	Papp – high
MTS = 50 μM	MTS = 200 μM	MTS = 225 μM
Easy chemistry	Hard chemistry	Difficult to patent

Compound 4	Compound 5	Compound 6
Ki = 10 μM	Ki = 0.05 μM	Ki = 0.2 μM
MW = 229	MW = 300	MW = 450
Papp – high	Papp – medium	Papp – medium
MTS = 250 μM	MTS = 150 μM	MTS = 0 μM
Easy chemistry	Unstable in buffer	Fast chemistry

Papp = apparent cell permeability. Assay used to determine the rate at which a compound passes through a cell membrane. High rate is good to get oral bioavailable compounds.
MTS = assay to determine if the compound has an unwanted effect in cell proliferation (cell toxicity). The higher the MTS the less effect in cell proliferation (less toxic)

Lead Identification

Ligand Efficiency

- **Binding Efficiency Index (BEI)**

  Ligand binding increases with molecular weight but at the same time this decreases solubility and increases metabolism. BEI is used to compare Ligand efficiency of compounds

  BEI= *pKi, pKd or pIC50 (mol/L) /MW (kDa)* *the higher the BEI the better the lead*

- **Surface Binding Efficiency Index (SEI)**

  Taking into account PSA since it is related to how good the compound will be absorbed in the gut and therefore related to %F. PSA > 100 Å² leads to a rapid fall-off in %F

  SEI= *pKi, pKd or pIC50/PSA* *the higher the SEI the better the lead*

%F = bioavailability: fraction of an oral dose that reaches the systemic circulation

Lead Identification

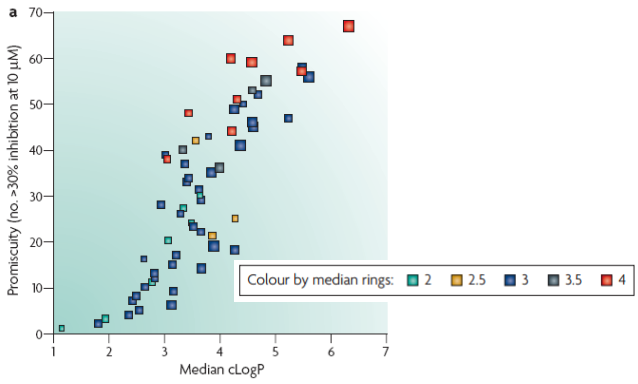
Choosing a Lead

Compound 1	Compound 2	Compound 3
Ki = 50 μM	Ki = 0.8 μM	Ki = 5 μM
MW = 410	MW = 350	MW = 250
Papp – very low	Papp – high	Papp – high
MTS = 50 μM	MTS = 200 μM	MTS = 225 μM
Easy chemistry	Hard chemistry	Difficult to patent
BEI = 10.5	BEI = 17.4	BEI = 21.2

Compound 4	Compound 5	Compound 6
Ki = 10 μM	Ki = 0.05 μM	Ki = 0.2 μM
MW = 229	MW = 300	MW = 450
Papp – high	Papp – medium	Papp – medium
MTS = 250 μM	MTS = 150 μM	MTS = 0 μM
Easy chemistry	Unstable in buffer	Fast chemistry
BEI = 21.8	BEI = 24.3	BEI = 14.9

Lead Identification

Ligand-Lipophilicity Efficiency



$LLE = pIC_{50} - clogP$

Important to avoid lipophilicity driven activity!

Nat Rev Drug Discov. 2007, 6, 881-890

Lead Optimization

Definition of terms

- **Lead Optimization**

  Is the synthetic modification of a biologically active compound to fulfil all physicochemical, pharmacokinetic, pharmacological and toxicologic requirements for clinical usefulness

- **Pharmacophore**

  an ensemble of steric and electronic features that is necessary for molecular recognition of a ligand by a biological macromolecule

Lead Optimization

Important considerations

- **Minimize pharmacophore**
- **Secure IP position**
- **Convergent synthetic routes where possible**
- **Selective/ Potent**
- **Optimize ADMET - Adsorption/Distribution/Metabolism/ Excretion/Toxicity**

  Plasma protein binding, solubility, metabolic stability, cell permeability, CYP inhibition/induction, hERG, selectivity, PK profile

- **Efficacy in relevant disease models**
- **Viable synthesis "Cost of Goods"**

Lead Optimization

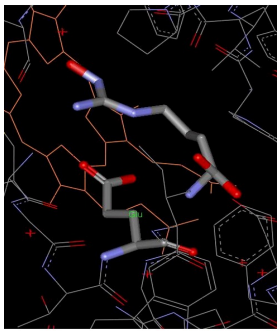
Computational Chemistry & Chemoinformatics

- Design Optimal "drug-like" Expansion Series
- Identify all building blocks available commercially
- Apply rule based filters - Lipinski, etc
- Build virtual library
- Dock with target if structure available
- Analyse diversity
- Generate databases
- Calculate electronic properties
- Statistical Analysis
- Rationalise large numbers of results
- QSAR models
- 3D Pharmacophore models
- ADMET models

Lead Optimization

Using structural data

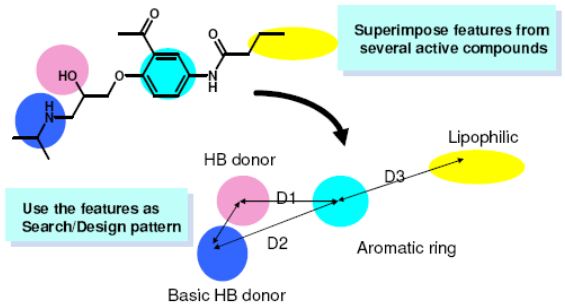
- Understanding how your compound interacts with the target protein
- Use this information to increase potency & selectivity



Lead Optimization

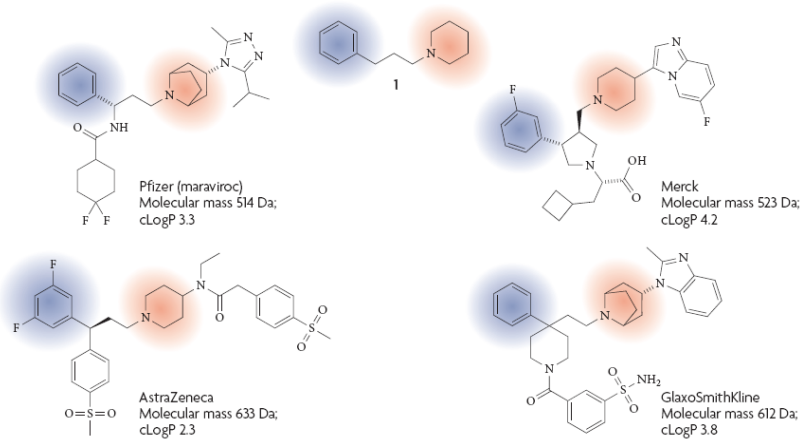
Pharmacophore Modelling

- Usually used when there is no protein structure
- Model is based on active and inactive compounds



Lead Optimization

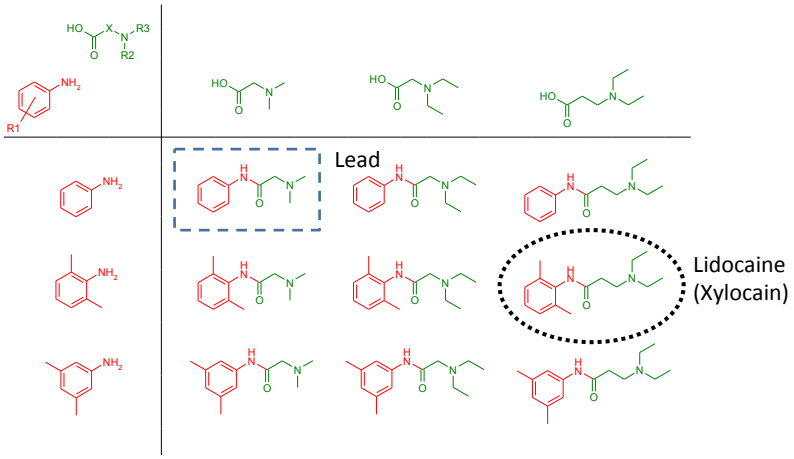
Pharmacophore Modelling



Nat Rev Drug Discov. 2007, 6, 881-890

Lead Optimization

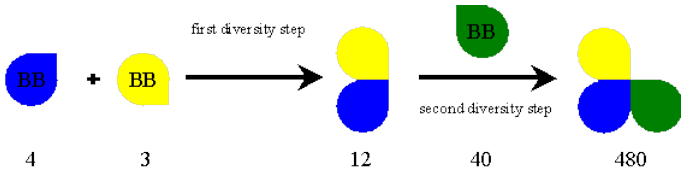
Library concept

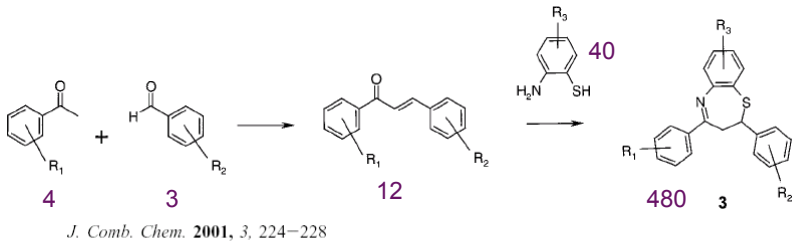


9 possible products from 6 building blocks

Lead Optimization

Multiple Parallel Synthesis (MPS)



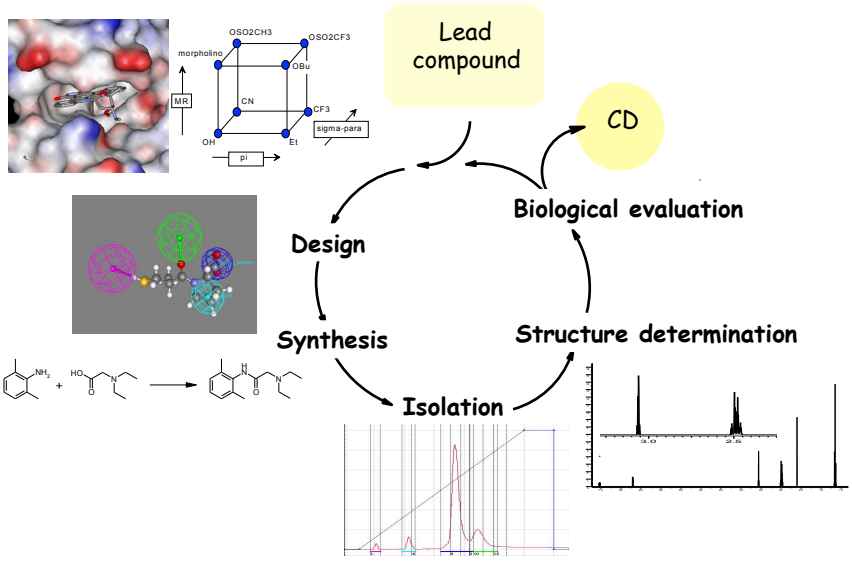


J. Comb. Chem. **2001**, *3*, 224−228

A medicinal chemist aims to make a few mg of many molecules in a short period of time

Lead Optimization

From Lead to CD



Lead Optimization

Metabolism and Toxicology

- **Pharmacokinetics**

  Study of absorption, distribution, metabolism and excretion (ADME) of bioactive compounds in a higher organism

- **Biotransformation**

  Is the chemical conversion of substances by living, organisms or enzyme preparations derived therefrom

- **Absorption**

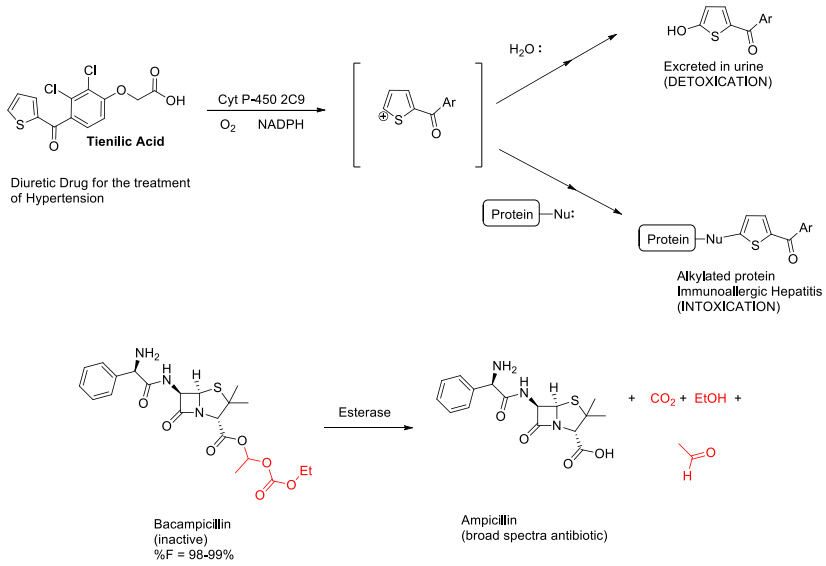
  Movement of a substance into the bloodstream

- **Metabolism**

  Chemical reactions involved in the maintenance and reproduction of life. Synthetic chemicals without a nutritive value (xenobiotics) are detoxified by metabolizing enzymes (mostly in the liver) normally by reducing lipophilicity in order to be excreted

Lead Optimization

Metabolism and Toxicology



Lead Optimization

Properties of a Candidate Drug

- **Novel structure**
  It must be patentable
- **Amenable to synthesis**
  Reasonable cost of production (industrial scale)
- **Good physical properties**
  Solubility, stability, etc
- **Active, selective and safe**
  Avoid off target activity to reduce side effects
- **Good pharmacokinetics**
  Bioavailability, distribution, metabolism, elimination

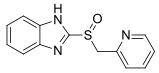
Lead Optimization

An example

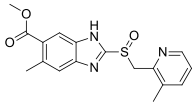
Lead – active
(side effect: inhibits iodine uptake)

Increased activity and selectivity – metabolism OK
(unwanted findings in dogs: vasculitis)

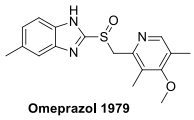
Optimized substance



Timoprazol 1973



Picoprazol 1976



Omeprazol 1979

Losec (thousands of substances had been synthesized – first approval 1987)

The role of the Medicinal Chemist